

CA
1
EP24
85
R17

CA 1 EP 2085R17

CANADIAN CLIMATE CENTRE

REPORT NO. 85-17

MONTHLY PREDICTION

BY THE

ANALOGUE METHOD

Earth Sciences Library

by

Amir Shabbar
and
John L. Knox

ATMOSPHERIC ENVIRONMENT SERVICE

DOWNSVIEW, ONTARIO

1985

THIS REPORT HAS RECEIVED ONLY LIMITED CIRCULATION BECAUSE OF THE SPECIAL NATURE OF THE PROJECT. HOWEVER, REFERENCE IS PERMITTED IF THE WORDS "UNPUBLISHED MANUSCRIPT" ARE MADE PART OF THE BIBLIOGRAPHIC ENTRY, IN ACCORDANCE WITH ACCEPTED PRACTICE.

OTHER CANADIAN CLIMATE CENTRE (CCC)
SCIENTIFIC SERIES INCLUDE:

1. CS series - CCC Climatological Study
2. CLI series - CCC Climatological Memoranda
3. CDS/DS series - CCC Climatic Data Summary
- CCC Documentation Sheet

For further information on climatological publications contact the

Climatological Services Division
Atmospheric Environment Service
4905 Dufferin Street
Downsview, Ontario M3H 5T4

MONTHLY PREDICTION
BY THE
ANALOGUE METHOD

by

Amir Shabbar
and
John L. Knox

A publication of the Canadian Climate Program



Digitized by the Internet Archive
in 2025 with funding from
University of Toronto

<https://archive.org/details/31761056885767>

TABLE OF CONTENTS

	Page
Summary	ii
1. Introduction	1
2. Objectives	1
3. Data Base	2
4. Selection Criteria	2
5. Methodology	3
6. Verification	4
7. Results	
7.1 Experiment 1 (Daily data)	4
7.2 Practical Ceiling of the Analogue Method	5
7.3 Ultimate Ceiling of the Analogue Method	5
7.4 Experiment 2 (5-day average data)	6
8. Discussion and Conclusions	7
9. References	8
10. Figures (1 - 9)	9

SUMMARY

The analogue technique was used to select anomalies of 50 kPa absolute (AT) and relative topography (RT) over the Northern Hemisphere from 20 to 180°W and from 30°N to the North Pole. Experiments were performed on the height anomalies of the daily and 5-day average data. For the selection criterion, both RMS and correlation techniques were investigated. A number of weighting systems were applied to the target sequence, including the 4th power law. For randomly chosen target sequences, verification of 50 kPa RT over North America shows the following: The spatial correlation technique appeared to give better results than the RMS. Statistically, objective analogue prediction is not as skillful as persistence using the last known instantaneous state of the atmosphere; on average, the top-ranked analogues verify the best and skill scores are similar for each of the four seasons. The best of all analogue in the archive outscored persistence by a wide margin from day 2 to 13. We also found that given a set of objectively selected top-ranking analogues, there exists, within that set, one member which significantly outscores persistence.

SOMMAIRE

La technique analogique sert à sélectionner des anomalies de topographie absolue (TA) et de topographie (TR) de 50 kPa au-dessus de l'hémisphère nord de 20 à 180°W et de 30°N au Pôle Nord. On a effectué des expériences sur les anomalies d'altitude des données moyennes quotidiennes et sur cinq jours. Pour la sélection des critères, on a examiné tant la technique de la moyenne quadratique que la technique de corrélation. On a appliqué plusieurs systèmes de pondération à la suite initiale, y compris la loi de la puissance 4. Pour les suites initiales choisies de façon arbitraire, la vérification de la TR de 50 kPa au-dessus de l'Amérique du Nord révèle ceci : La technique de corrélation spatiale a paru donner de meilleurs résultats que la moyenne quadratique. Sur le plan statistique, la prévision analogique objective n'est pas si juste que la persistance recourant au dernier état instantané connu de l'atmosphère; en moyenne, la série analogique présentant la meilleure corrélation est la mieux vérifiée et les indices de vérification sont analogues pour chacune des quatre saisons. La meilleure série analogique des archives a de beaucoup dépassé, en justesse, la persistance du 2^e jour au 13^e. Nous avons aussi constaté que dans une série analogique de données présentant la meilleure corrélation et sélectionnées de façon objective, il existe un élément qui, en justesse, dépasse nettement la persistance.

MONTHLY PREDICTION BY THE ANALOGUE METHOD

by

Amir Shabbar and John L. Knox

1. INTRODUCTION

In March, 1985 the Atmospheric Environment Service of Canada undertook (for the first time in its history) the issue of monthly and seasonal forecasts of temperature and precipitation anomalies on an experimental basis.

In earlier decades (e.g. Namias, 1951) reservations were expressed concerning the viability of using analogues for extended-range forecasting. This was understandable in view of the relatively limited archive of hemispheric upper air data at that time. Lorenz (1969) attempted to investigate atmospheric predictability by selecting the 2 best matching analogues of the 3-dimensional Northern Hemisphere continuum. Hopefully, they would be so similar that, for prediction purposes, they could be regarded as identical initial states. Unfortunately the investigation was confined to only 5 years of data and so, not surprisingly, the search was unsuccessful. More recently, Barnett and Preisendorfer (1978) addressing the problem of seasonal forecasts, chose analogues on the basis of "climate state vector" similarity, and these in turn were used to predict seasonal surface air temperature, and precipitation anomalies for N. America. It should also be noted that since 1976, the Canadian Meteorological Centre at Montréal has been issuing 15-day forecasts using analogue methods as described by Knox & Robinson (1983).

The principle of the particular analogue technique which we shall use, is to identify those past atmospheric states which most closely resemble the current state and its evolution during the past "n" days. It is hoped that the subsequent development of the current state, at least for a limited period of time, will be similar to that of the best analogue(s) found in the archives. The advantage of predictions made in this way is that they automatically contain the climatology of the locations for which the forecast is required. They also provide sample scenarios showing how the forecast mean might be constituted. We do not expect that analogues selected using our "matching of the current evolving state" technique (to be described later) will have significant predictive skill beyond 10 to 15 days. If, however such skill is attainable then, by selecting analogues on the basis of a sequence of suitably weighted numerical weather prediction (NWP) prognoses out to say 10 days, it should be possible to incorporate both NWP and analogue predictive skill into the preparation of a forecast out to 20 to 25 days.

2. OBJECTIVES

Specifically our objectives are:

- a) to choose the spatial and temporal domain of the data at CMC (Knox and Robinson, 1983) so as to optimize the analogue selection;
- b) to best define "current evolving state" by examining different weighting system;

- c) to test statistical techniques for analogue selection, e.g. correlation (COR), root mean square (RMS), and possibly covariance (COV);
- d) to compare analogue prediction skill with simply persisting the current day into the forecast period;
- e) to compare the best a posteriori analogue (i.e. the one which gave the best forecast) with the a priori objectively chosen top-ranking analogue;
- f) to compare the "ultimate" with the "practical" (i.e. attainable) predictive limits of the analogue methodology;
- g) to investigate if, by subjective intervention, one can improve the analogue skill score.

3. DATA BASE

The basic archive consisted of once-daily (1200Z) data for the Northern Hemisphere at 2 levels, 500mb and MSL, from January 1, 1946 to October 31, 1984, or nearly 39 years of record. After converting the MSL data set into 1000mb height, daily hemispheric fields at 500mb R.T. were prepared. The analogue selection technique was then based on contemporaneous "n" day sequences of data at two pressure surfaces:

- a) 500mb Absolute Topography (A.T.)
- b) 500mb Relative Topography (R.T.) i.e. 500mb-1000mb thickness

After considerable experimentation we chose a spatial domain of the 500mb A.T. which (Figure 1) encompassed a sector of the N.H. bounded on the west by 180°, on the east by 20°W, on the south by 30°N, and on the north by the Pole (a total of 161 lat.-long. intersections on the Figure 1 grid). For the 500mb R.T., the domain was restricted to North America (as approximated by the 91 lat.-long. intersections indicated in Figure 1). The reason for so confining the 500mb R.T. domain is the high frequency of occurrence of equivalent barotropic structures in amplified height fields over the oceans. In such areas, 500mb R.T. anomalies have a high correlation with 500mb A.T. anomalies and therefore do not add significant information for extended range forecast purposes. On the other hand, the baroclinicity arising from ocean-continent thermal contrasts gives rise, particularly during the winter season, to thickness anomalies often out-of-phase with 500mb anomalies. Therefore both anomaly fields are taken into consideration over the North American continent for analogue selection purposes. We admit that the above choice of spatial domain may not be optimum.

4. SELECTION CRITERIA

Clearly it is important to develop appropriate criteria for measuring the degree of similarity of relevant features between 2 atmospheric fields. In a study of Wintertime Analogs of the 500mb Height field, Gutzler and Shukla (1984) did parallel selection runs using respectively, the 3 following measures of similarity:

- a) minimization of Root Mean Square difference (RMS);
- b) maximization of Spatial Correlation (COR);
- c) maximization of Spatial Covariance (COV).

It is interesting to note that the corresponding lists of 10 ranked cases by no means showed a one-to-one correspondence, and indeed, there were few commonalities between the RMS list on the one hand and the COR and COV lists on the other. The reason seems to be that RMS and COR criteria place different emphasis on features of spatial similarity. The RMS metric is very sensitive to amplitude and phase differences between strong systems and is relatively insensitive to large phase differences between weak amplitude systems. On the other hand, COR maximization ensures sensitivity to phase correspondence between systems but is completely insensitive to their relative amplitudes.

We carried out a large number of comparisons between the two measures. For example in Figure 2(a) we show the 500mb anomaly target field (5-day mean August 6-10, 1981). The COR method chose September 13-17, 1967, Figure 2(b) as the second-ranking analogue, while the RMS method ranked that same field as a very low 193! It is obvious that these charts do have fairly close correspondence over the spatial domain to which the criteria were applied (E. Pacific, N. America, and N. Atlantic). Many other cases gave similar results and so we decided that, notwithstanding its aforementioned shortcoming, COR appeared to be a better measure.

5. METHODOLOGY

The target, or "current evolving state" is defined as an n -day sequence ($n \leq 10$) of the atmospheric states specified by the 500mb A.T. and R.T. anomaly fields for the respective domains. The current day is designated by k . The search for the best analogue sequences is confined to a window of 100 days centred on k . In other words we attempt to find analogues "within season". In effect, for any given target, there are ninety 10-day sequences to be examined for each year of record, or a total of 3,510 for the 39-year archive. To avoid redundancy, it was stipulated that analogues must not overlap by more than 5 days, the rationale for which will become evident shortly.

The weighting system was designed to emphasize the current trend. Slug-gish response to rapid large scale changes will result if the weights are assigned equally to each of the individual days in the target. At the other extreme total loading on the current day was found to result in excessive sensitivity to highly amplified but transient development. As a compromise between these two extremes, a fourth power weighting system was introduced (Figure 3). From the curve it is clear that over 90% of the weight is on the front 5 days of the sequence.

The following is a summary of the analogue selection process.

- a) For the target sequence ending on current day " k ", calculate the spatial correlations $r_k, r_{k-1}, \dots$ etc, between members of the 161 point 500mb A.T. sequence and corresponding members of the potential analogue sequences.
- b) Repeat the procedure for the 500mb R.T. fields.

- c) In order to mimic 3 dimensionality in the model, average these 2 correlations for each day of the sequence.

e.g.
$$R_k = 1/2 [(r_k)_{A.T.} + (r_k)_{R.T.}]$$

- d) Using the 4th power weighting system take the weighted average over n days ($n \leq 10$).

$$C = \sum_{k=0}^9 w_k R_k$$

- e) Rank the analogues according to the values of C. List the date, the correlation R_k for each day of the sequence and the C values.

6. VERIFICATION

Because of the strong relationship between thickness and surface temperature, we shall confine the presentation of verification results to analogue predictions of 500mb R.T. anomalies over N. America. We calculated, for each of the subsequent 30-days, spatial correlations between the actual (observed) field and the analogue prediction for the 91-point domain.

To measure analogue skill we used persistence. The field for the current day was simply persisted for each of the 30-days and corresponding spatial correlations were calculated. This produced a forward stepping autocorrelation time series. Verification results will be presented and discussed in the next section.

7. RESULTS

7.1 Experiment 1

For 60 randomly chosen targets throughout the archives (1946-1984) and for each season the skill of the average top 10 ranking analogues was assessed relative to persistence. The verification was carried out by calculating the correlation coefficient (for the corresponding daily values) between the predicted sequence and the verifying sequence. The results of Experiment 1 (Figure 4) show that on the average; persistence outscores analogue in each of the 4 seasons. The analogue skill decays at the same rate as that of persistence. Student's t-test shows that the analogue scores are significantly different from 0 for the first 3 days (a period where persistence scores are higher); beyond day 6 in the verification period, both analogue and persistence scores are poor.

7.2 Practical Ceiling of the Analogue Method

If a forecaster always selected the best verifying analogue among those objectively chosen, how well would he do? This establishes a practical ceiling of the analogue method. In order to find this ceiling we modified the procedure of Experiment 1 so that instead of verifying the average of the top 10 ranking analogues we carried out verification on the best verifying analogues in the top 10 ranks. The best verifying analogue was identified as that analogue which has the largest correlation coefficient averaged over the 10-day verifying sequence. Figure 5 shows the verification scores (average of 60 cases) for each of the four seasons. For comparison, persistence of the last day in the target sequence is also plotted. For each of the four seasons, the best verifying analogue in the top 10 out performs persistence from Day 2 to about Day 12. However at the 5% significance level, the correlation coefficients are not different from zero after Day 10. The analogue scores are significantly different from that of persistence from Day 5 to 8 for Autumn, from Day 7 to 10 for Winter and for Days 7 and 8 for Summer. For Spring, the analogue method still outscores persistence; however, the score are not as large as in the other seasons and they barely make the statistical significance criteria. These results suggest that indeed if we had some means, e.g. enlightened synoptic judgement, of picking out the best verifying analogue in the top 10 ranks as selected by the object procedure, then on the average that particular analogue would out perform persistence from Day 5 to Day 10. In our long-range prediction program, we have been using synoptic judgement to choose analogue(s) from the objectively selected ones.

7.3 Ultimate Ceiling of the Analogue Method

One would like to find out on the basis of a posteriori verification (highest correlation coefficient averaged over 10 days) what is the best analogue that is available in the archives. This upper limit is termed "Ultimate Ceiling" or "best of all". If we had a technique for discovering such analogues, then our results would verify at the "Ultimate Ceiling" level.

The magnitude of the verifying correlation coefficient of the "best of all" analogue determines the potential of the analogue technique. The average scores of the "best of all" analogues are plotted against persistence for each of the four seasons in Figure 6. These analogues outscore persistence from Day 2 to about Day 13, and the two scores are statistically distinguishable at the 5% level. Note that despite good scores in the prediction period, the *a priori* correlation coefficients with the target sequence are very low. The magnitude of the verifying correlation coefficient of the "best of all" analogue determines the potential of the analogue technique.

For 60 randomly chosen targets, the distribution of the "best of all" analogue over the thickness domain for each of the four seasons is shown in Figure 7. The numbers in Figure 7 set the upper limit of success using analogue selection. For most of the winter cases, the correlation coefficient falls between 0.50 and 0.55. For the Spring season, the peak of the

distribution is broader and occurs between 0.40 and 0.55. During Summer and Autumn, the distribution peak at lower correlations. The highest correlation value (0.72) occurs in winter. Barnett and Preisendorfer (1978) also noted that their highest "global best" values occurred in the winter season.

How does the top ranking analogue compare with the ultimate ceiling level? In Figure 8 we attempt to answer this question. The ratio of the correlation of the top ranking analogue to the "best of all" is about 0.30. This ratio is slightly higher for the Winter and Spring seasons; however, large standard error (≈ 0.3) does not permit the acceptance of the results as statistically distinguishable. How often and in which ranks does the "best of all" analogue occur? The numbers, in Table 1, represent the ranks of the analogue (as chosen by the selection criterion) for which the best verification occurred. Only the top forty ranks were considered. There is no clear pattern, the best verifying analogue seems to occur evenly through the forty ranks.

7.4 Experiment 2

Filtering the data by forming 5-day means would eliminate highly transient synoptic features. We suspect that there will be better continuity and fewer cases of top-ranking analogues suddenly disappearing from the list the following day. Since we are concerned with finding the closest fit between a single 5-day mean anomaly and the single 5-day mean anomalies in the archives, the calculations are much simpler. Potential analogues are searched within the season of the target and no overlap is permitted. For the same random targets as in Experiment 1, the skill of the average top 10 ranking analogues was assessed relative to the persistence of the target 5-day mean. Figure 5 shows the verifying correlation coefficient (over the thickness domain) by matching 5-day means in the prediction period. For all seasons, persistence outcores analogue. Winter scores are slightly better than the other seasons. The analogue scores are statistically significant at 5% level at the first 5-day in the prediction period for all seasons. The scores are indistinguishable from zero by day 10.

Table 1 The numbers represent the ranks of the analogue for which the "best of all" analogue occurred amongst the top 40 ranks.

Winter	Spring
2, 4, 5, 7, 10, 14, 15, 17, 34, 40	3, 4, 8, 10, 19, 20, 24, 28, 29
Summer	Autumn
1, 3, 8, 9, 14, 15, 20, 22, 23, 26, 32, 40	3, 4, 5, 6, 9, 10, 21, 24, 27, 29, 38

8. DISCUSSION AND CONCLUSIONS

We found that for analogue selection, the spatial correlation technique appeared to give better results than the RMS. However, from a review of the literature, we were unable to find evidence that a systematic comparative research into the various techniques for objective selection of analogues of large scale atmospheric circulations has ever been carried out. This is an area closely related to the more general problem of pattern recognition where as yet there is "no powerful theory ready to come to the aid of any particular application" (Somerville, 1976). Principal component analysis, frequently used in map-typing investigations, may well have a useful application for analogue selection. We propose to test this technique in a subsequent study.

Our results appear to indicate that, on average, the so-called "top" analogues, based on an objective ranking system, have less predictive skill than persistence.

By using verification after the fact, we investigated the "ultimate" limits of analogue predictability. Such analogues outscored persistence by a wide margin from Day 2 to Day 13. However, their correlation with the target sequence was on average quite low, and so they seldom appeared in the top ranking list. This suggests that such analogues often occur as a result of substantial readjustment in the large-scale circulation modes during the prediction period. It is conceivable that some of these might be revealed if analogue selection were based on NWP fields (out to Day 5 at least) instead of on current and past data. In general however, we are not at all sanguine that it will be possible to achieve the "ultimate" limit of analogue predictability.

We also investigated the more practical question of whether, given a set of objectively selected top-ranking analogues, there exists, within that set, one member which significantly outscores persistence. This was found to be the case, and our conclusion is that it might be possible to have improved selection mechanism. Enlightened synoptic judgement may very well be one such method. We hope to substantiate this statement in a later study, however, for over 2 years the authors have examined top-ranked sets of analogues of 500mb A.T. issued once daily in real-time from CMC, Montréal. It is our distinct impression that by a systematic synoptic evaluation, one can usually select one or two analogues which will score higher than the average of the group.

9. ACKNOWLEDGEMENT

The authors wish to express their thanks to Drs. S.W. Woronko and G.J. Boer for their suggestions and review of the manuscript. The authors appreciate the assistance of Mr. T. Jang in establishing the computer program and running the various phases of calculations. Ms. U. Ellis deserves credit for typing the manuscript.

REFERENCES

- Barnett, T.P., and R.W. Preisendorfer, 1978: Multifield Analog Prediction of Short-term Climate Fluctuations using a Climate State Vector. J. Atmos. Sci., 35, 1771-1787.
- Gutzler, D.S., and J. Shukla, 1984: Analogs in the Wintertime 500 mb Height Field. J. Atmos. Sci., 41, 177-189.
- Knox, J.L. and R.H. Robinson, 1983: Extended Range Forecasts for Canada using Analogues of the 500mb Height Field. Proceedings of the Eighth Annual Climate Diagnostics Workshop, U.S. Dept. of Commerce, NOAA.
- Lorenz, E.N., 1969: Atmospheric Predictability as Revealed by Naturally Occurring Analogs. J. Atmos. Sci., 26, 636-646.
- Namias, J., 1951: General Aspects of Extended Range Forecasting. Compendium of Meteorology, American Meteorological Society, 802-813.
- Somerville, R.C.J., 1976: Pattern Recognition and Multivariate Verification. Weather Forecasting and Weather Forecasts, NCAR Colloquium: Summer 1976: Vol. 2, 776-794.

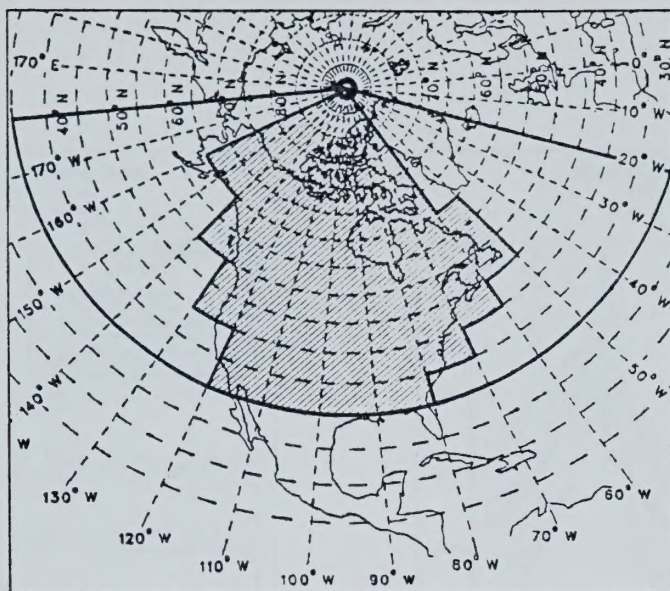


Figure 1 CMC archive grid. The domain of the 500mb height anomaly is delineated by heavy line. Shaded area shows the 500-1000mb thickness anomaly domain.

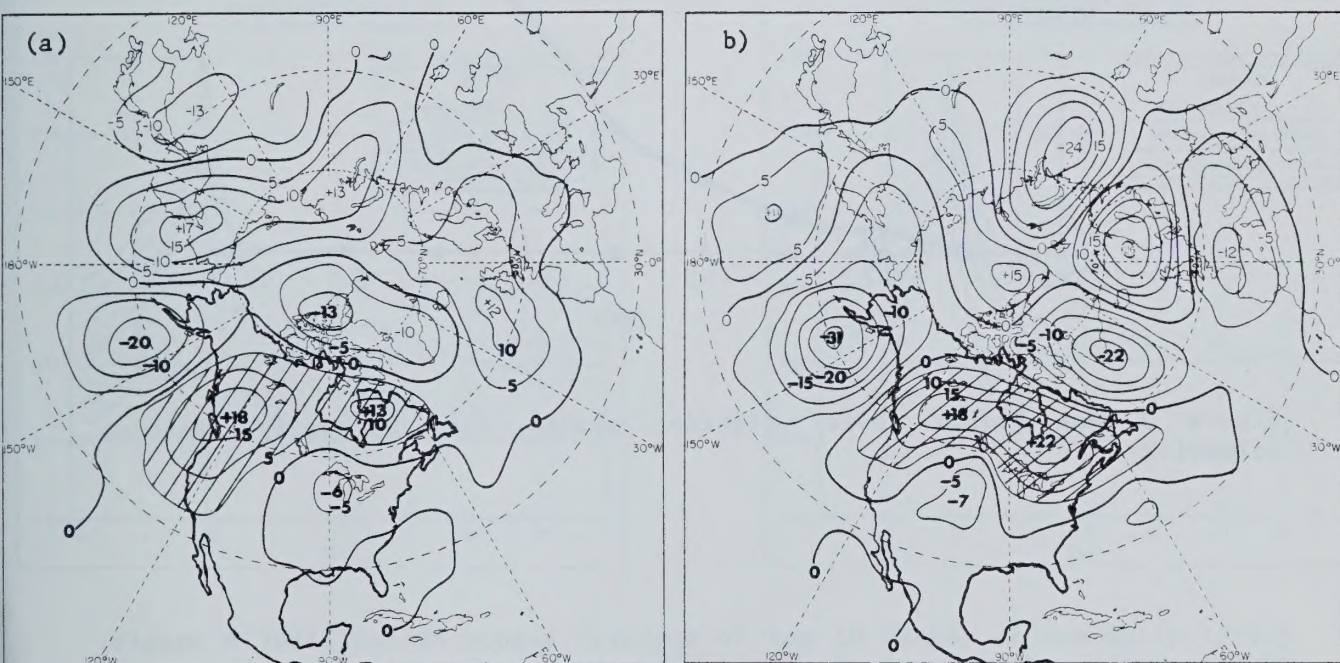


Figure 2 (a) 5-day average of 500 mb anomaly target ending on August 10, 1981; b) analogue of (a) ending on September 17, 1967 which ranked 2 by the correlation method and 193 by the RMS method.

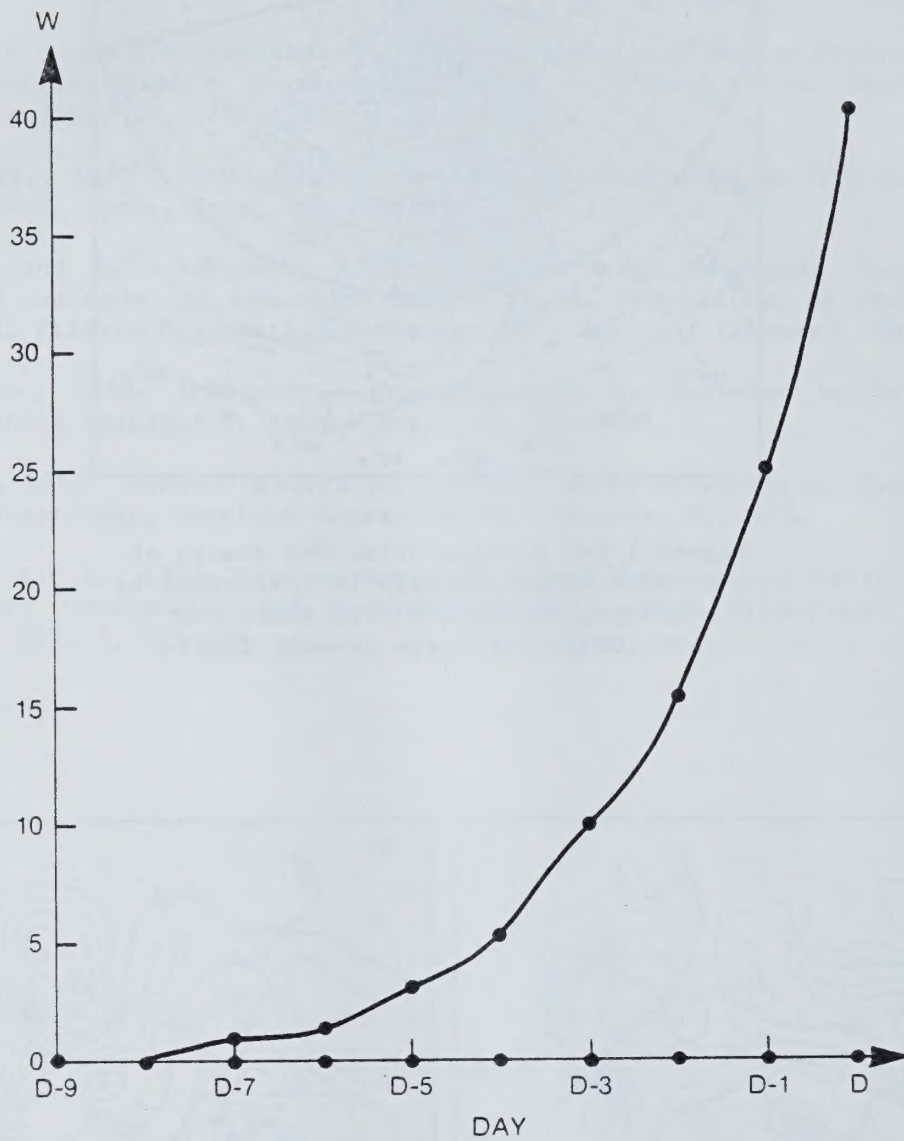


Figure 3 The fourth power weighting system for the daily analogue selection.

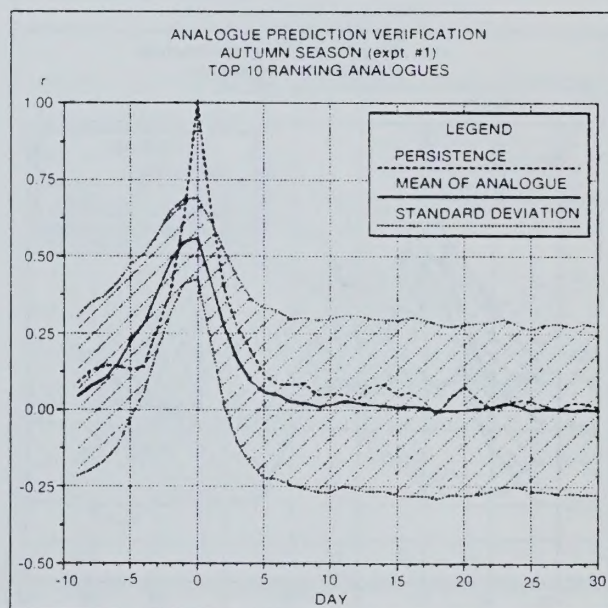
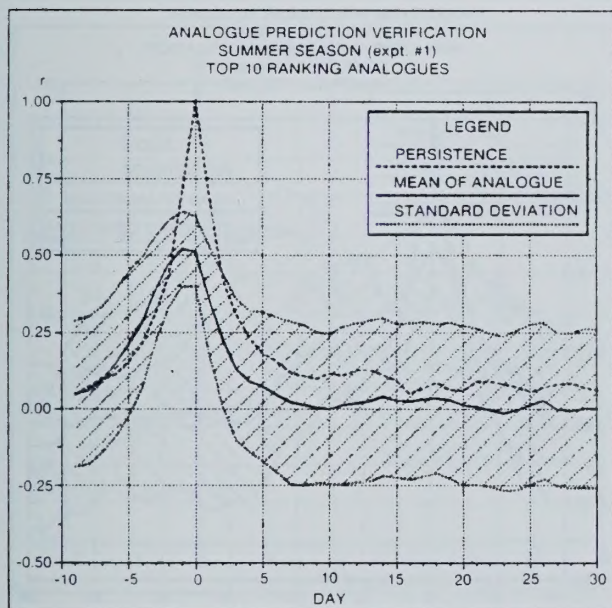
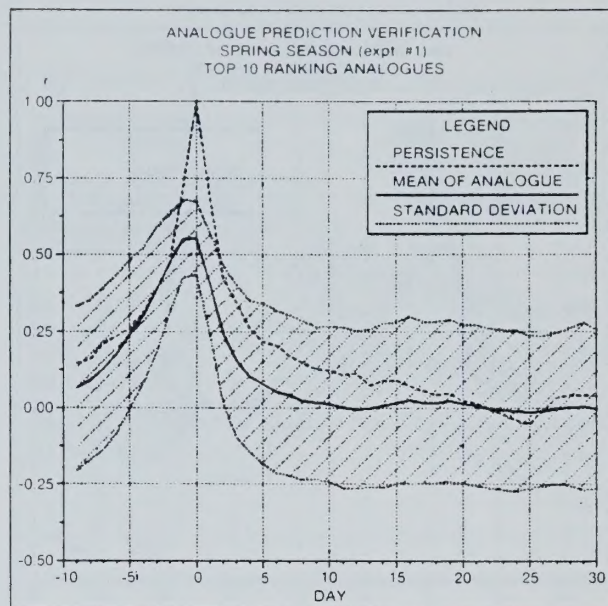
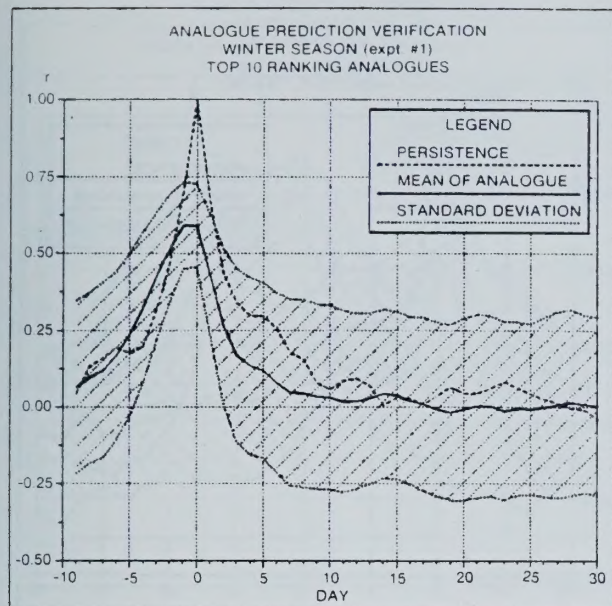


Figure 4 Verification scores (average of top 10 ranks) of the daily 10-day sequence target plotted against persistence of the last day in the target sequence. An average of 60 random targets were considered for each of the four seasons.

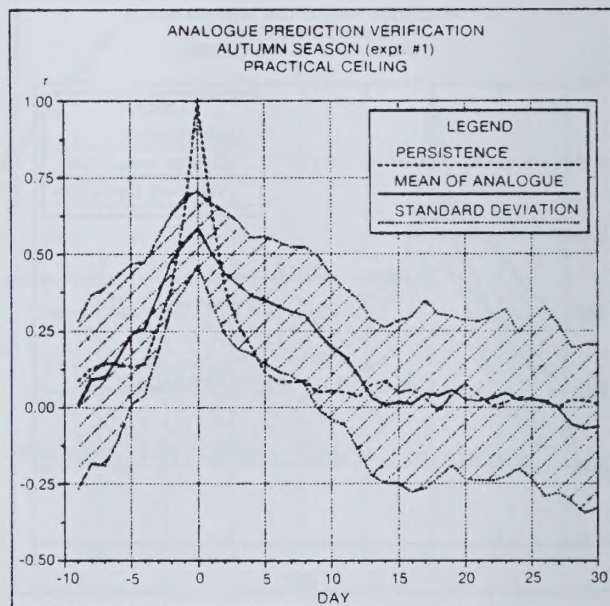
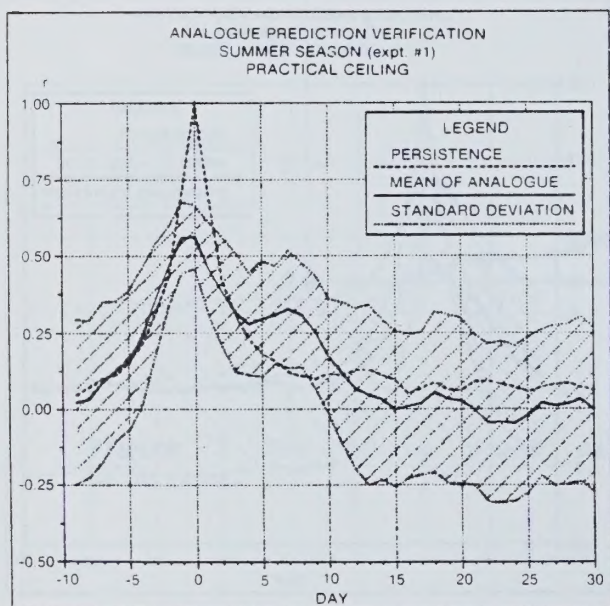
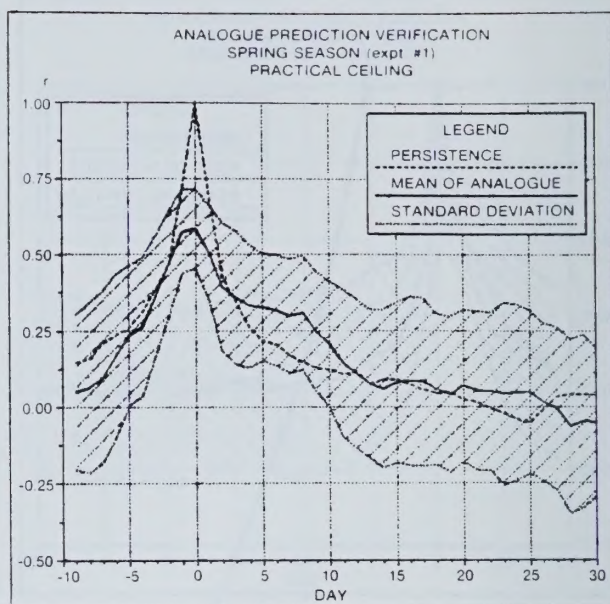
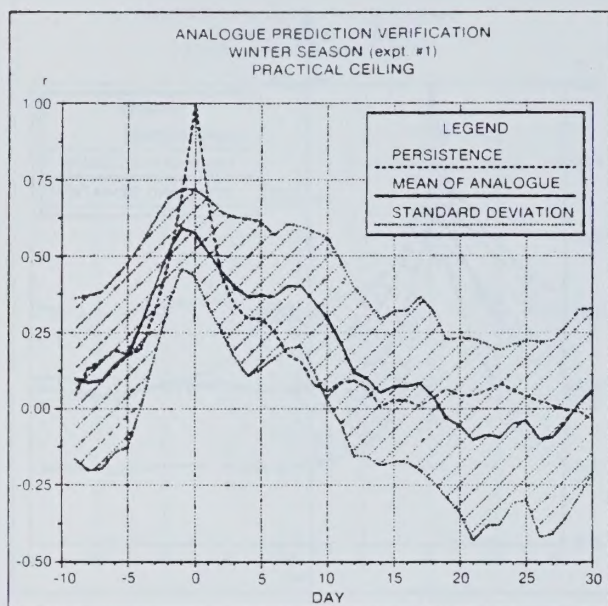


Figure 5 Verification scores of the best verifying analogue in the top 10 ranks plotted against persistence of the last day in the target sequence. An average of 60 random targets were considered for each of the four seasons.

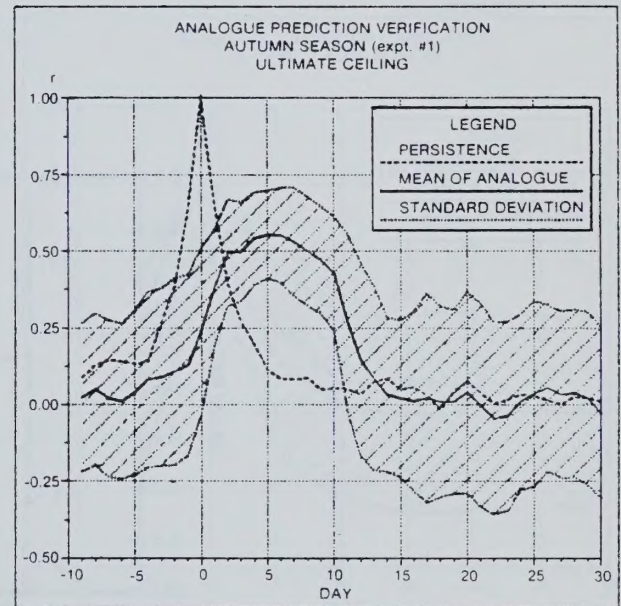
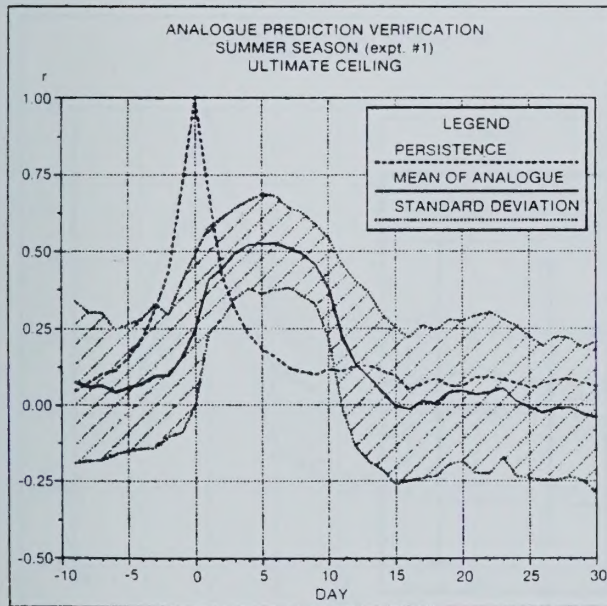
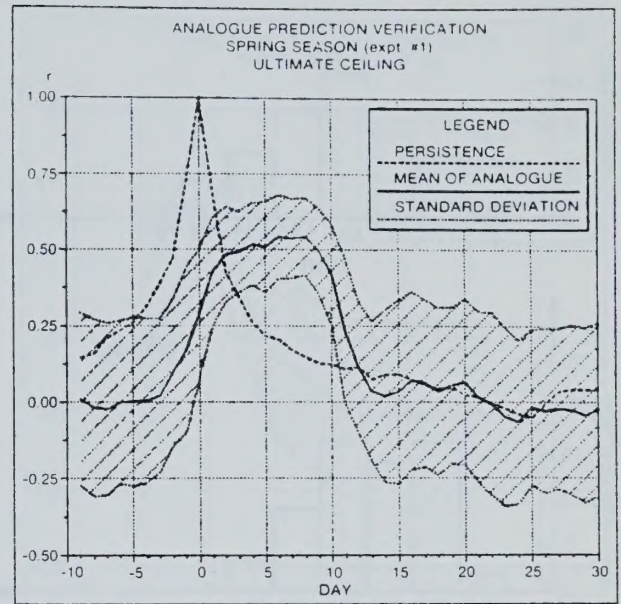
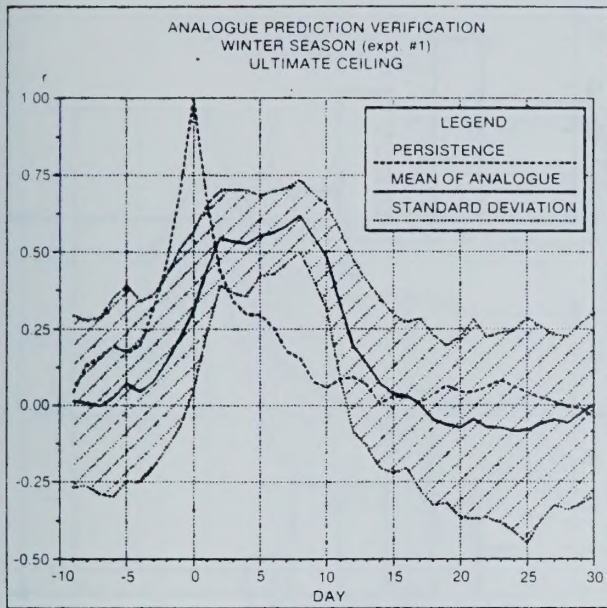


Figure 6 Verification scores of the "best of all" analogues plotted against persistence of the last day in the target sequence. An average of 60 random targets were considered for each of the four seasons.

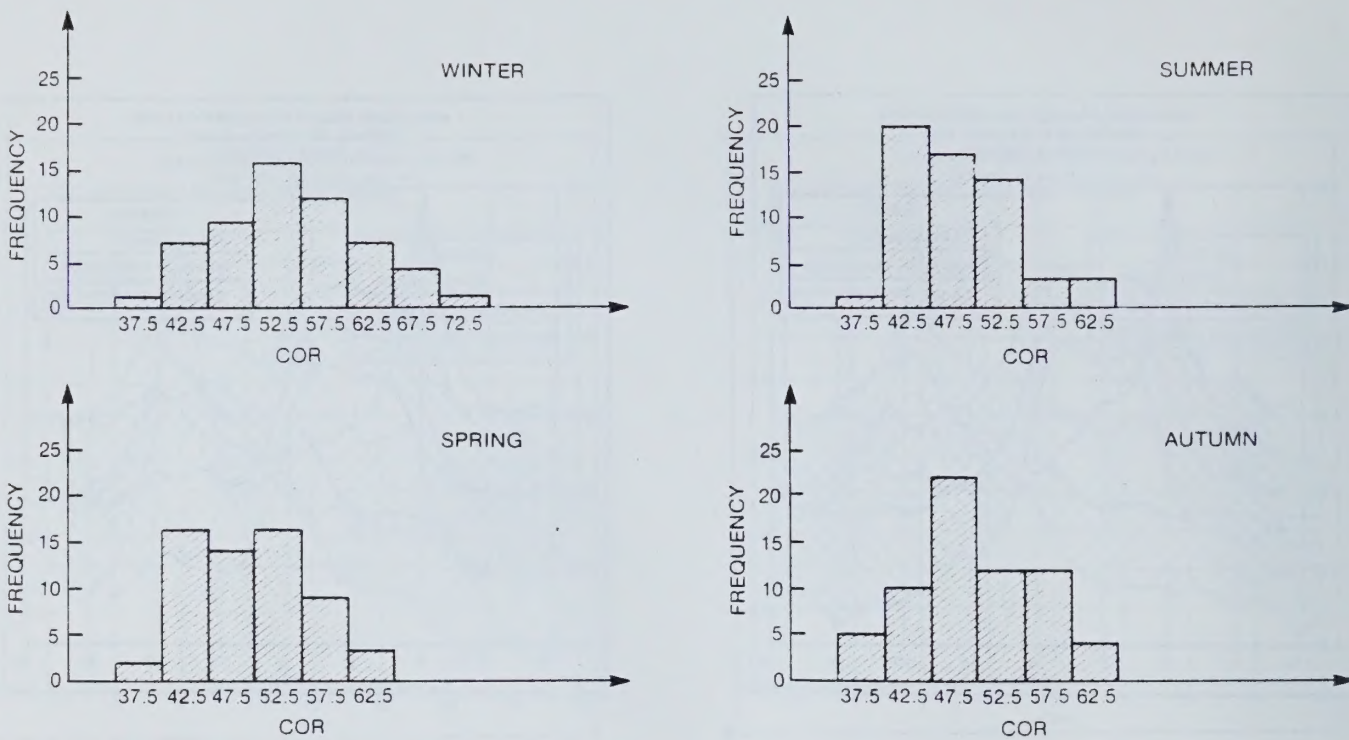


Figure 7 Bar chart of the "best of all analogue in the archives for each season.

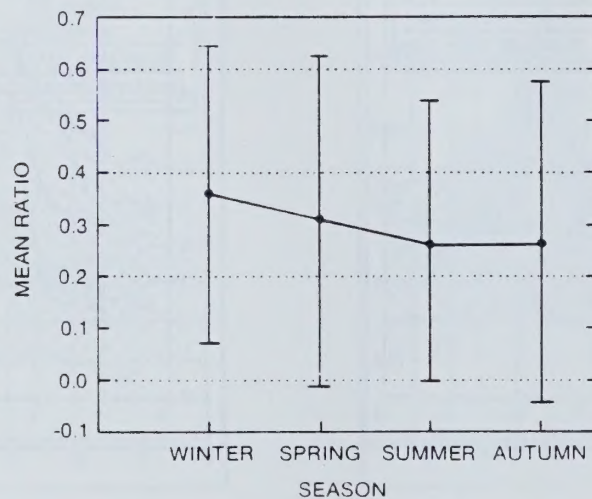


Figure 8 Ratio of the correlation of the best analogue chosen by the selection criterion to that of the "best of all" found in the archives.

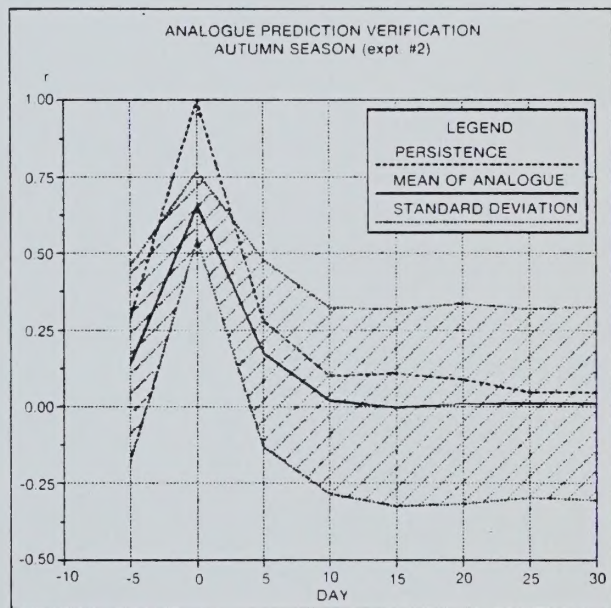
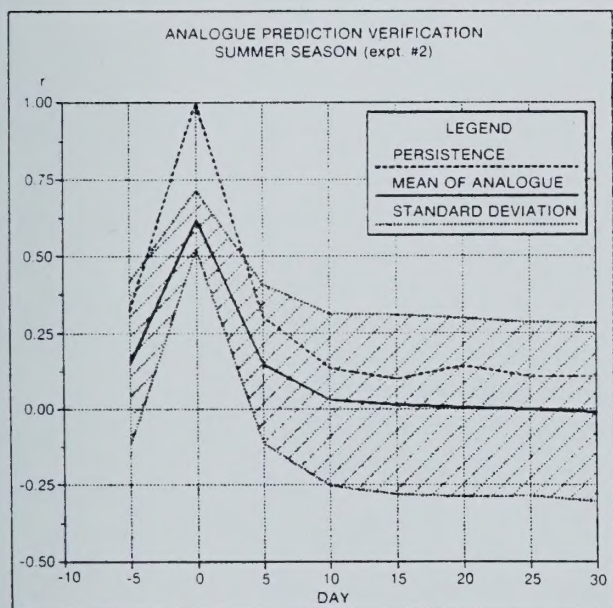
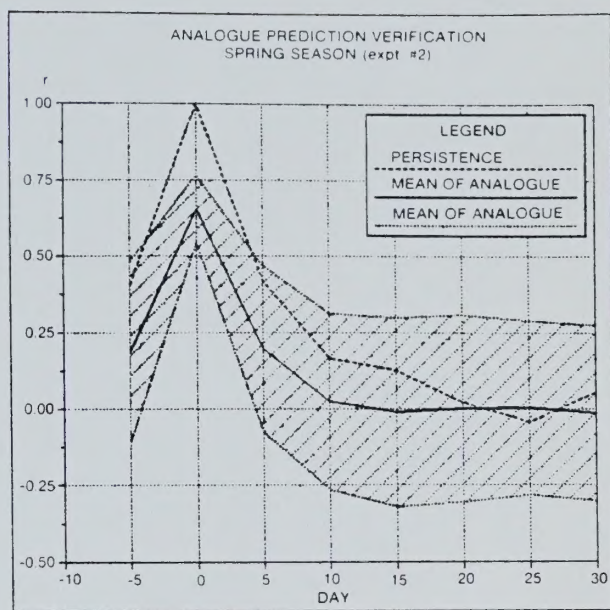
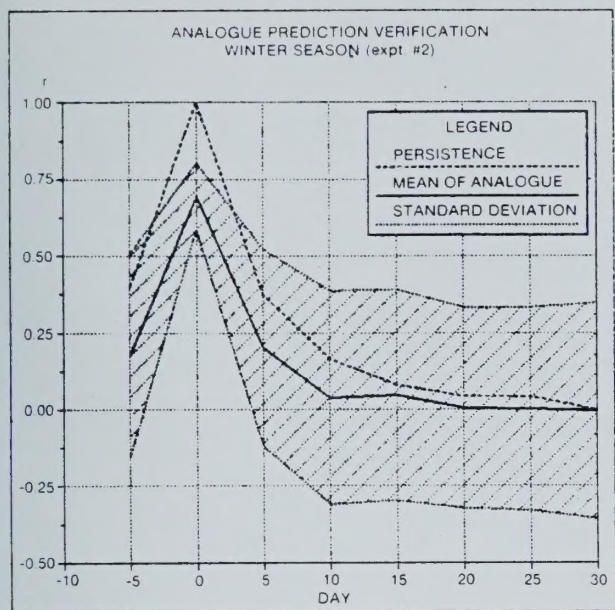


Figure 9 Verification scores (average of top 10 ranks) of the 5-day mean target sequence plotted against persistence of the target 5-day mean. An average of 60 random targets were considered for each of the four seasons.

